Chapter 3

Case 2: Capital Asset Pricing Model

Overview

In this project, you will investigate the capital asset pricing model (CAPM). You will compute the IBM's beta exposure to the market. You need to download the dataset "Case1CAPM.csv" on blackboard and program in R studio.

Data Description

- DATE : Date formed as year month and day. For example, 3July2017 is entered as 20170703.
- IBMRET: **Daily** International Business Machines (IBM) stock returns **in percentage unit** defined in equation 2.1.
- MarketEXRET: **Daily market excess returns in percentage unit** defined as $R_{m,t} - r_{f,t}$. The excess returns on the market, value-weight returns of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ that have a CRSP share code of 10 or 11.
- RF: The Tbill return is the simple **daily rate (in percentage)** that over the number of trading days in the month compounds to 1-month TBill rate from Ibbotson and Associates Inc.
- Data Source: CRSP through WRDS and Fama French Data Library.

3.1 Introduction

Modern financial theory rests on two assumptions: (1) securities markets are very competitive and efficient (that is, relevant information about the companies is quickly and universally distributed and absorbed); (2) these markets are dominated by rational, risk-averse investors, who seek to maximize satisfaction from returns on their investments.

Although these two assumptions constitute the cornerstones of modern financial theory, the formal development of CAPM involves other, more specialized limiting assumptions. These include frictionless markets without imperfections like transaction costs, taxes, and restrictions on borrowing and short selling. The model also requires limiting assumptions concerning the statistical nature of securities returns and investors preferences. Finally, investors are assumed to agree on the likely performance and risk of securities, based on a common time horizon.

3.1.1 Diversification

An underpinning of CAPM is the observation that risky stocks can be combined so that the combination (the portfolio) is less risky than any of its components. Although such diversification is a familiar notion, it may be worthwhile to review the manner in which diversification reduces risk.

Some of the risk assumed is peculiar to the individual stocks in their portfolios — for example, a companys earnings may plummet because of a wildcat strike. On the other hand, because stock prices and returns move to some extent in tandem, even investors holding widely diversified portfolios are exposed to the risk inherent in the overall performance of the stock market.

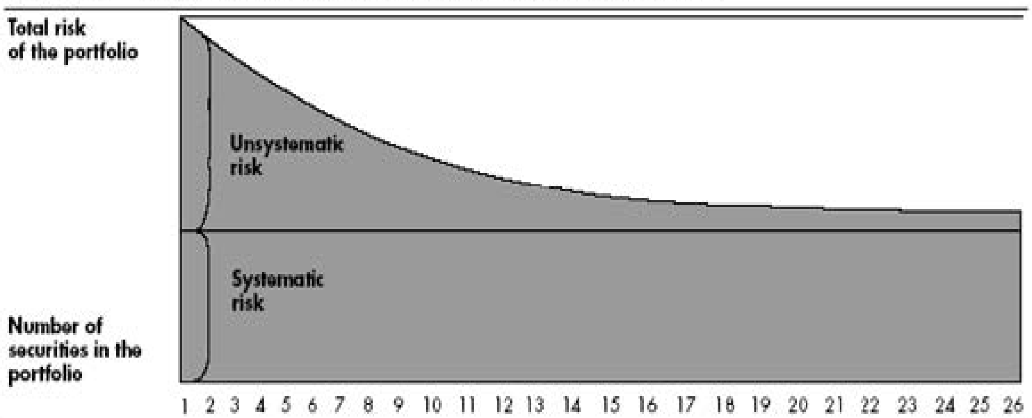
So we can divide a securitys total risk into unsystematic risk, the portion peculiar to the company that can be diversified away, and systematic risk, the nondiversifiable portion that is related to the movement of the stock market and is therefore unavoidable. Examples of systematic and unsystematic risk factors appear in Exhibit I.

Exhibit I **Some unsystematic and systematic risk factors**

Unsystematic risk factors
A company's technical wizard is killed in an auto accident.
Revolution in a foreign country halts shipments of an important product ingredient.
A lower-cost foreign competitor unexpectedly enters a company's product market.
Oil is discovered on a company's property.
Systematic risk factors
Oil-producing countries institute a boycott.
Congress votes a massive tax cut.
The Federal Reserve steps up its restrictive monetary policy.
Long-term interest rates rise precipitously.

Exhibit II graphically illustrates the reduction of risk as securities are added to a portfolio. Empirical studies have demonstrated that unsystematic risk can be virtually eliminated in portfolios of 30 to 40 randomly selected stocks. Of course, if investments are made in closely related industries, more securities are required to eradicate unsystematic risk.

Exhibit II **Reduction of unsystematic risk through diversification**



Citation:
1:Does the Capital Asset Pricing Model Work? by David W. Mullins
2: CampbellLoMacKinley1997

3.1.2 Theory: Security Market Line

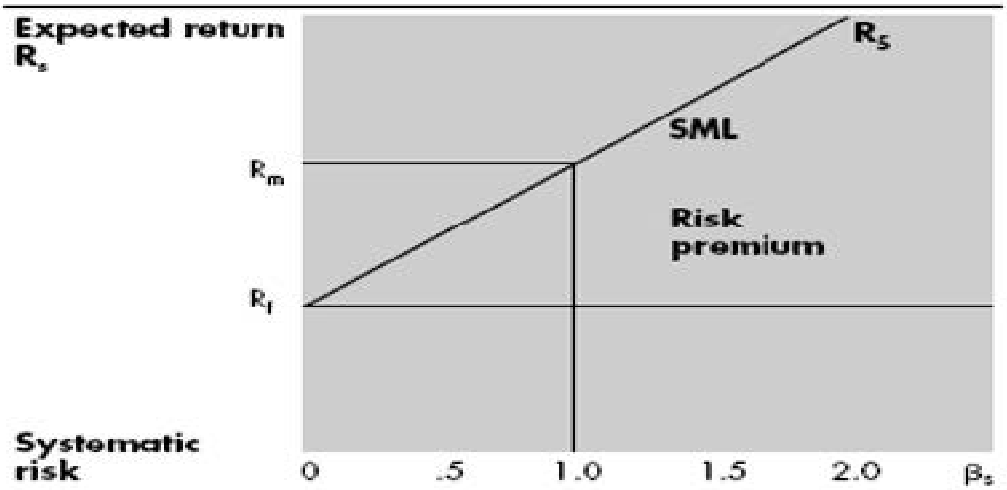
The culmination of the sequence of conceptual building blocks is CAPMs risk/expected return relationship. This fundamental result follows from the proposition that only systematic risk, measured by beta β, matters. Securities are priced such that:

$$E_t(R_{i,t+1}) = R_{f,t} + \beta_i(E_t(R_{m,t+1}) - R_{f,t}) \tag{3.1}$$

where: $E_t(R_{i,t+1})$ = the stocks expected return (and the companys cost of equity capital); $R_{f,t}$ = the risk-free rate from t to t+1, t refers to the time when you invest; in the lecture we define the same risk free as $r_{f,t+1}$, t+1 refers to the time when you cash out; $E_t(R_{m,t+1})$ = the expected return on the stock market as a whole and β_i = the stocks beta = $\frac{Cov_t(R_{i,t+1}, R_{m,t+1})}{Var_t(R_{m,t+1})}$.

This risk/expected return relationship is called the security market line (SML)(graphically illustrated in Exhibit III).

Exhibit III The security market line



Where does it come from? From the above discussion of capital market line and tangency portfolio, we know that nobody can do better than holding the market. The tangency portfolio must be the market. Suppose that for two assets A and B: $\frac{E_t(R_{A,t+1}) - R_{f,t}}{Cov_t(R_{A,t+1}, R_{M,t+1})} > \frac{E_t(R_{B,t+1}) - R_{f,t}}{Cov_t(R_{B,t+1}, R_{m,t+1})}$ (Asset A offers a better return/risk ratio than asset B), then buy A, sell B. What if everybody does this? Hence, in equilibrium, all return/risk ratios must be equal for all assets, then it is the one of the market: $\frac{E_t(R_{A,t+1}) - R_{f,t}}{Cov_t(R_{A,t+1}, R_{m,t+1})} > \frac{E_t(R_{B,t+1}) - R_{f,t}}{Cov_t(R_{B,t+1}, R_{m,t+1})} =$

$\frac{E_t(R_{m,t+1})-R_{f,t}}{Var_t(R_{m,t+1})}$. This gives the relationship between risk and expected return for individual stocks and portfolios.

The expected return on a security generally equals the risk-free rate plus a risk premium. In CAPM the risk premium is measured as beta (number of units of risk) times the expected return on the market minus the risk-free rate (risk premium per Unit). The risk premium of a security is a function of the risk premium on the market, $E_t(R_{m,t+1}) - R_{f,t}$, and varies directly with the level of beta. No measure of unsystematic risk appears in the risk premium, of course, for in the world of CAPM diversification has eliminated it.

Beta is the standard CAPM measure of systematic risk. It gauges the tendency of the return of a security to move in parallel with the return of the stock market as a whole. One way to think of beta is as a gauge of a securitys volatility (of the market-wide shock) relative to the markets volatility. A stock with a beta of 1 an average level of systematic risk rises and falls at the same percentage as a broad market index, such as Standard & Poors 500-stock index. Stocks with a beta greater than 1.00 tend to rise and fall by a greater percentage than the market — that is, they have a high level of systematic risk and are very sensitive to market changes. Conversely, a stock with a beta less than 1.00 has a low level of systematic risk and is less sensitive to market swings.

In the freely competitive financial markets described by CAPM, no security can sell for long at prices low enough to yield more than its appropriate return on the SML. The security would then be very attractive compared with other securities of similar risk, and investors would bid its price up until its expected return fell to the appropriate position on the SML. Conversely, investors would sell off any stock selling at a price high enough to put its expected return below its appropriate position. The resulting reduction in price would continue until the stocks expected return rose to the level justified by its systematic risk.

One perhaps counterintuitive aspect of CAPM involves a stock exhibiting great total risk but very little systematic risk. An example might be a company in the very chancy business of exploring for precious metals. Viewed in isolation the company would appear very risky, but most of its total risk is unsystematic and can be diversified away. The well-diversified CAPM investor would view the stock as a low-risk security. In the SML the stocks low beta would lead to a low risk premium. Despite the stocks high level of total risk, the market would price it to yield a low expected return.

In summary, if the model correctly describes market behavior, the relevant measure of a securitys risk is its market-related, or systematic, risk measured by beta. If a securitys return bears a strong positive relationship with the return on the market and thus has a high beta, it will be priced to yield a high expected return; if it has a low beta, it will be priced to yield a low expected return.

Since unsystematic risk can be easily eliminated through diversification, it does not increase a securitys expected return. According to the model, financial markets care only about systematic risk and price securities such that expected returns lie along the security market line.

Citation:

1:Does the Capital Asset Pricing Model Work? by David W. Mullins

2: Global Financial Management lecture notes by Campbell Harvey and Stephen Gray.

3: Efficiently Inefficient by Lasse H. Pedersen

3.2 Analysis

Step 1: Load data Case1CAPM.csv

```
#Case 2: CAPM (Chap9)
CAPM<-read.csv("Case1CAPM.csv", header = TRUE, sep=",")
```

Step 2: Change the format of variable "DATE"

```
DATE<-as.Date(as.character(CAPM$DATE), "%Y%m%d")
```

Step 3: Create the excess returns of IBM

```
ibmRET<-CAPM$IBMRET
marketEXERT<-CAPM$MarketEXRET
RF<-CAPM$RF
IBMEXERT<-ibmRET-RF
```

3.2.1 Estimate Single Index Model

Step 4: Model Setup

RW

In the single index model, we derived a relation between the excess returns on an individual security and the market excess returns. We can write this ordinary least squares regression:

$$R_{i,t+1} - R_{f,t} = \underbrace{\alpha_i}_{\text{Nonmarket Premium}} + \underbrace{\beta_i}_{\textit{Sensitivity}} (R_{m,t+1} - R_{f,t}) + \underbrace{\epsilon_{i,t+1}}_{\text{Diversifiable Surprises}} \tag{3.2}$$

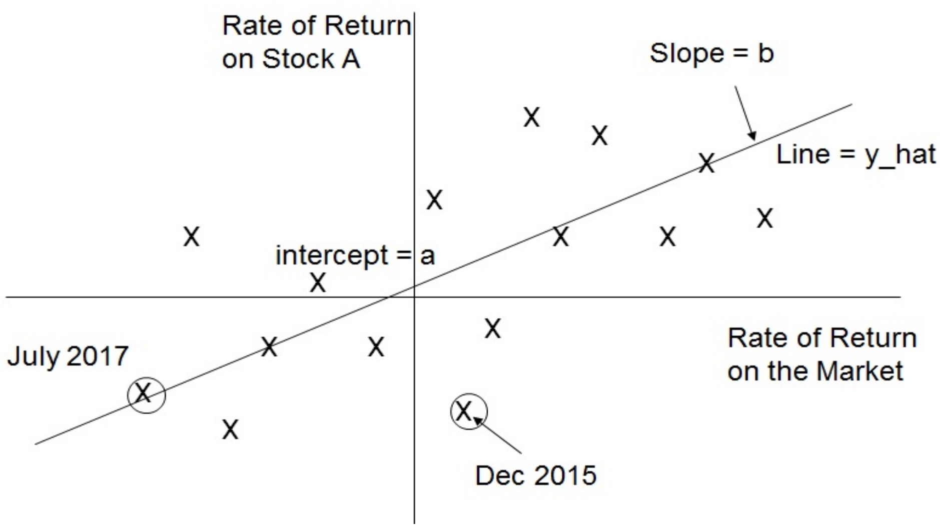
This holds for all i. The alpha is the intercept in the regression, which is the non-market premium. The beta is the covariance between the security i's returns and the market returns divided by the variance of the market returns and $E_t\epsilon_{i,t+1} = 0$. The beta risk is referred to in some text books as the sensitivity to the systematic or non-diversifiable or market movement. This risk is rewarded with expected returns. There is another type of risk which is called non-systematic or diversifiable, nonmarket or idiosyncratic risk, presented by the term $\epsilon_{i,t+1}$.

When fitting a least squares line, we generally require:
a: Linearity. The data should show a linear trend.
b: Nearly normal disturbance: Idiosyncratic Risk $\epsilon_{i,t+1} \sim N(0, \sigma_t^2)$
c: Constant variability $Var_t(\epsilon_{i,t+1}) = \sigma_t^2 = \sigma^2$
d: Independent observations: Idiosyncratic Risk $\epsilon_{i,t+1}$ is independent of its past and its future observations — $\epsilon_{i,t+1}$ is serially independent.

Step 5: Model Estimation

EST

The asset's characteristic line is the line of the best fit for the scatter plot that represents simultaneous excess returns on the asset and on the market.



This is just the fitted values from a regression line. The b will be the regression slope (estimated β) and the a will be the intercept (estimated α). The error in the regression is the distance from the line (predicted) to each point on the graph (actual).

Let $y_{i,t+1} = R_{i,t+1} - R_{f,t}$, $X_{i,t+1} = R_{m,t+1} - R_{f,t}$ and the residuals are the leftover variation in the data after accounting for the model fit,

$$\widehat{y_{i,t+1}} = \underbrace{a_i + b_i X_{i,t+1}}_{\text{Fitted Value}}; \qquad \underbrace{e_{i,t+1}}_{\text{Residuals}} = y_{i,t+1} - \widehat{y_{i,t+1}} \tag{3.3}$$

A hat on y is used to signify that this is an estimate. This estimate may be viewed as an average: the equation predicts that stock i will have an average return of $a_i + b_i * 10\%$, when the market excess return is 10% ($X_{i,t+1} = 10\%$). Absent further information about other market and firm conditions, the prediction for the expected excess returns of stock i that uses the average is a reasonable estimate.

Each observation will have a residual. If an observation is above the regression line, then its residual, the vertical distance from the observation to the line, is positive. Observations below the line have negative residuals. One goal in picking the right linear model is for these residuals to be as small as possible.

How to determine the best line? Fitting linear models by eye is open to criticism since it is based on an individual preference. In this case, we use least squares regression as a more rigorous approach — choose the line that minimizes the sum of the squared residuals,

$$e_{i,1}^2 + ... + e_{i,T}^2 \tag{3.4}$$

The line (a_i and b_i) that minimizes this least squares criterion is represented as the solid line in the above figure. This is commonly called the least squares line. The following are three possible reasons to choose Criterion 3.4

a. It is the most commonly used method.

b. Computing the line based on Criterion 3.4 is easy by hand and in most statistical software.

c. In many applications, a residual twice as large as another residual is more than twice as bad. Squaring the residuals accounts for this discrepancy.

```
Model<-lm(IBMEXERT~marketEXERT)
```

lm(y∼x) constructs linear regression model. y is the dependent variable and x is the independent variable. By default, intercept is included.

```
## get the coefficient panel
Model$coefficients
## get acess to intercept and slope
Model$coefficients[1]
Model$coefficients[2]
```

Estimated Regression Line

$$\widehat{IBMexRTN} = \underset{(0.019613)}{0.04042} + \underset{(0.016525)}{0.91177} \times MarketexRTN \tag{3.5}$$

Step 6: Model Interpretation

RW

Interpreting parameters in a regression model is often one of the most important steps in the analysis. Suppose that he slope (b_i) and intercept(a_i) estimates are $b_i = 1.5$ and $a_i = 3\%$. What do these numbers really mean?

The slope describes the estimated difference in the y variable if the explanatory variable x for a case happened to be one unit larger. The intercept describes the average outcome of y if $x = 0$ and the linear model is valid all the way to $x = 0$, which in many applications is not the case.

Interpreting the slope parameter is helpful in almost any application. For each additional 1% increase in the market excess returns, we would expect the stock i to receive a net increase of $1.5 * 1\% = 1.5\%$ in its excess returns on average. Note that a higher market excess returns corresponds to higher excess returns of stock i because the coefficient is positive.

The estimated intercept $a_i = 3\%$ describes the excess returns of stock i if a the market excess returns is zero. The meaning of the intercept is relevant to this application since the market excess returns can be zero at some point of time. In other applications, the intercept may have little or no practical value if there are no observations where x is near zero.

```
Modelsum<-summary(Model)
Modelsum
```

summary(Mymodel) can output the important statistics for linear regression model **Mymodel**. It will output statistics such as R square, beta coefficients and p-value.

```
Call:
lm(formula = IBMEXERT ~ marketEXERT)

Residuals:
     Min       1Q   Median       3Q      Max 
-13.8015  -0.6594  -0.0582   0.5996  12.9184 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.04042    0.01961   2.061   0.0394 *  
marketEXERT  0.91177    0.01653  55.175   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.459 on 5538 degrees of freedom
Multiple R-squared:  0.3547,	Adjusted R-squared:  0.3546 
F-statistic:  3044 on 1 and 5538 DF,  p-value: < 2.2e-16
```

Formula Call As you can see, the first item shown in the output is the formula R used to fit the data. Note the simplicity in the syntax: the formula just needs the predictor (market excess returns) and the target/response variable (IBM excess returns).

```
## get the coefficient panel
Modelsum$coefficients
```

Coefficients The third section in the model output talks about the coefficients of the model. Theoretically, in simple linear regression, the coefficients are two unknown constants that represent the intercept α and slope terms β in the linear model. Ultimately, the analyst wants to find an intercept a and a slope b such that the resulting fitted line is as close as possible to the data points in the data set.

Coefficient - Estimate The coefficient Estimate contains two rows: the first row is the intercept; the second row is the slope.

- **Intercept 0.04**: The estimated daily risk adjusted excess return of IBM is 0.04% on average.
- **Slope 0.91**: For each additional 1% increase in the daily market excess return, we would expect an increase of $0.91 * 1\% = 0.91\%$ in the daily excess return of IBM on average. The sensitivity beta is estimated to be 0.91.

Coefficient - Standard Error The coefficient Standard Error measures the average amount that the coefficient estimates vary from the actual average value of our response variable. We'd ideally want a lower number relative to its coefficients. The Standard Error can be used to compute an estimate of the expected difference in case we ran the model again and again.

- **Intercept SD 0.0196**: The predicted daily risk adjusted excess return of IBM is 0.04% on average. This predicted average can vary by 0.0196%.
- **Slope SD 0.0165**: For each additional 1% increase in the daily market excess return, we would expect an increase of $0.91 * 1\% = 0.91\%$ in the daily excess return of IBM on average. This estimated sensitivity can vary by 0.0165.

Coefficient - t value The coefficient t-value is a measure of how many standard deviations our coefficient estimate is far away from 0. In our example, the t-statistic value for slope is very far away from zero which could indicate a relationship between IBM excess returns and market excess returns exists. The t-statistic value for intercept is also far away from zero which could indicate a nonzero risk adjusted excess return on average. In general, t-values are also used to compute p-values.

Coefficient - p-value The $Pr(t > |tstats|)$ acronym found in the model output relates to the probability of observing any value equal or larger than t statistic in absolute value. A small p-value for slope indicates that it is unlikely we will observe a relationship between the predictor (market excess returns) and response variables (IBM excess return) due to chance. A small p-value for intercept indicates that it is unlikely we will have a nonzero risk adjusted excess return on average.

Typically, a p-value of 5% or less is a good cut-off point. In our model example, the p-values are close to zero. Note the "signif. Codes" associated to each estimate. Three stars (or asterisks) represent a highly significant p-value.

Step 7: Goodness of Fit

RW

We evaluated the strength of the linear relationship between two variables earlier using the correlation, R. However, it is more common to explain the strength of a linear fit using R^2, called R-squared. If provided with a linear model, we might like to describe how closely the data cluster around the linear fit.

The R^2 of a linear model describes the amount of variation in the response that is explained by the least squares line. For example, the variance of the response variable, stock i's excess returns $y_{i,t+1} = R_{i,t+1} - R_{f,t}$, is $TSS = 29.8$. However, if we apply our least squares line, then this model reduces our uncertainty in predicting excess returns using the market excess returns. The variability in the residuals describes how much variation remains after using the model: $ESS = 22.4$. In short, there was a reduction of

$$R^2 = \frac{TSS - ESS}{TSS} = \frac{29.822.4}{29.8} = \frac{7.5}{29.8} = 0.205 \tag{3.6}$$

or about 25% in the datas variation by using information about market excess returns for predicting stock i's excess returns using a linear model. This corresponds exactly to the R-squared value: $R = 0.499$ or $R^2 = 0.25$.

Multiple R-squared, Adjusted R-squared The R-squared (R^2) statistic provides a measure of how well the model is fitting the actual data. It takes the form of a proportion of variance. R^2 always lies between 0 and 1 (i.e.: a number near 0 represents a regression that does not explain the variance in the response variable well and a number close to 1 does explain the observed variance in the response variable).

In our example, the R^2 we get is 0.3547. Or roughly 35% of the variance found in the response variable (IBM excess returns) can be explained by the predictor variable (market excess returns).

A side note: In multiple regression settings, the R^2 will always increase as more variables are included in the model. Thats why the adjusted R^2 is the preferred measure as it adjusts for the number of variables considered.

```
mdlpred<-predict(Model)
mdlresid<-resid(Model)
```

predict() is a generic function for predictions from the results of various model fitting functions.

resid() is a generic function which extracts model residuals from objects returned by modeling functions

Residuals The second item in the model output talks about the residuals. Residuals are essentially the difference between the actual observed response values (IBM excess returns) and the response values that the model predicted. The Residuals section of the model output breaks it down into 5 summary points. When assessing how well the model fit the data, you should look for a symmetrical distribution across these points on the mean value zero (0). If we see that the distribution of the residuals do not appear to be strongly symmetrical (a symmetric distribution will have 1Q+3Q close to zero, median close to zero and min+max close to zero), it means that the model predicts certain points that fall far away from the actual observed points.

Residual Standard Error Residual Standard Error is measure of the quality of a linear regression fit. Theoretically, every linear model is assumed to contain an error term ϵ. Due to the presence of this error term, we are not capable of perfectly predicting our response variable from the predictor one. The Residual Standard Error is the average amount that the response will deviate from the regression line.

In our example, the actual IBM excess returns can deviate from the regression line by approximately 1.459%, on average.

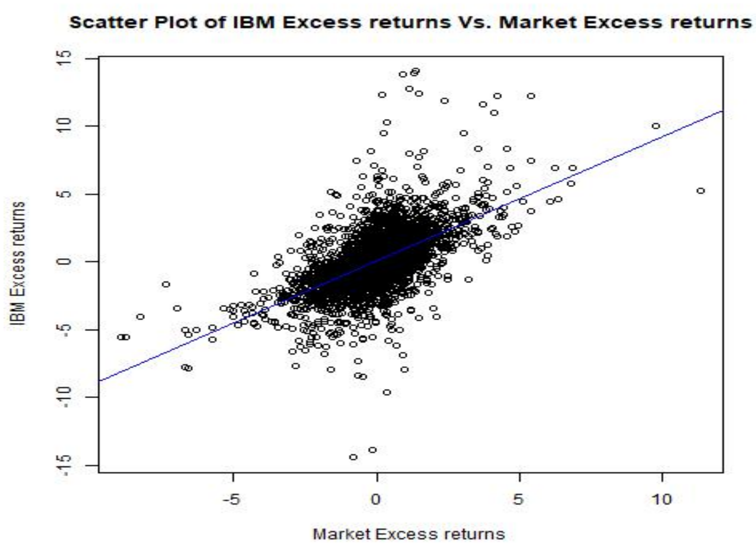
Its also worth noting that the Residual Standard Error was calculated with 5538 degrees of freedom. Simplistically, degrees of freedom are the number of data points that went into the estimation of the parameters used after taking into account these parameters (restriction). In our case, we had 5540 data points and two parameters (intercept and slope).

Step 8: Plot the OLS Line

FIG

```
jpeg(filename = "Case1_OLSLINE.jpeg")
plot(marketEXERT, IBMEXERT,
main="Scatter Plot of IBM Excess returns Vs. Market Excess returns",
xlab= "Market Excess returns", ylab="IBM Excess returns")
abline(Model, col="blue")
dev.off()
```

abline(..) This function adds one or more straight lines through the current plot.



3.2.2 Testing CAPM

The capital asset pricing model (CAPM) imposes the following constraint on expected returns for all securities i:

$$\alpha_i = 0 \tag{3.7}$$

The security's expected excess returns is linear in the security's beta. The beta represents the risk of security i in the market portfolio – or the contribution of security i to the variance of the market portfolio. The beta risk is the only type of risk that is rewarded or priced in equilibrium. What makes the CAPM different from the statistical model is that the CAPM imposes the constraint that the intercept or alpha is zero.

$$\alpha_i = \underbrace{\alpha_i + \beta_i(E_t(R_{m,t+1}) - R_{f,t})}_{\text{Actual Performances}} - \underbrace{\beta_i(E_t(R_{m,t+1}) - R_{f,t})}_{\text{Expected Performance}} \tag{3.8}$$

Thus, the alpha measure risk adjusted performance of a security.

One test of the CAPM is to test whether the alpha of any security or portfolio is statistically different from zero. The regression would be run with available stock returns data. The null hypothesis is (the CAPM holds) is that the intercept is equal to zero. Under the alternative hypothesis, the intercept or alpha is not equal to zero. The standard test is a t-test on the intercept of the regression. If the intercept is more than 2 standard errors from zero (or having a t-statistic greater than 2), then there is evidence against the null hypothesis (the CAPM).

a:The beta may not be constant through time. $\beta_{i,t} = \frac{Cov_t(R_{i,t+1}, R_{m,t+1})}{Var_t(R_{m,t+1})}$. We will discuss about this in Case 3.

b:The alpha may not be constant through time. If for some period, $\alpha_i > 0$, and others $\alpha_i < 0$. Even though we can't reject that alpha is zero, it doesn't necessarily mean that CAPM holds — risk adjusted returns is not zero. We will discuss about this in Case 3.

c: The errors may be correlated through time (this is known as autocorrelation or serial correlation).

d: Returns may be non-linearly related to market returns rather than the linear relation that is suggested in the statistical model. See discussion of derivatives in the textbook .

e: There may be other sources of risk.

Step 9: Hypothesis Test: t-stats

In this section we discuss uncertainty in the estimates of the slope and y-intercept for a regression line. We first discuss standard errors for these new estimates using statistical software and then we will perform hypothesis test.

Step 9.1: HT Is the risk adjusted return zero? To asses the validity of this claim, we run the above linear model estimation: the excess returns of stock i ($y_{i,t+1} = R_{i,t+1} - R_{f,t}$) against the market excess returns ($X_{i,t+1} = R_{m,t+1} - R_{f,t}$) with an intercept α_i.

Suppose that a_i is positive. This y-intercept is only the estimate of the parameter values α_i. We might wonder, is this convincing evidence that the true linear model has a zero intercept? That is, do the data provide strong evidence that CAPM is accurate? We can frame this investigation into a two-sided statistical hypothesis test:

$$\begin{aligned} H_0 : \quad & \alpha_i = 0. && \text{The risk adjusted excess return is zero.} \\ H_A : \quad & \alpha_i \neq 0. && \text{The risk adjusted excess return is not zero.} \\ & && \text{It could either be positive or negative.} \end{aligned}$$

We would reject H_0 in favor of H_A if the data provide strong evidence that CAPM holds. To assess the hypotheses, we identify a standard error for the estimate a_i, compute an appropriate test statistic, and identify the p-value.

In the hypotheses we consider, the null value for the intercept is 0, so we can compute the test statistic using the t (or z) score formula:

$$t - stats = \frac{\text{estimate} - \quad \text{null value}}{\text{Standard Error}} \tag{3.9}$$

We can look for the one-sided p-value using the probabilty table for the t-distribution. Or we could have identified the t-test statistic from the software output, shown in the first row and third column (t value). The entry in the first row and last column represents the p-value for the two-sided hypothesis test where the null value is zero. The corresponding one-sided test would have a p-value half of the listed value. Most of the time, we set the level of significance to be 5%.

$$\begin{aligned} p - value < 5\% \quad & \Longleftrightarrow |t - stats| > t_{df, 5\%/2} \\ p - value \geq 5\% \quad & \Longleftrightarrow |t - stats| < t_{df, 5\%/2} \end{aligned}$$

If $p - value < 5\%$, there is an evident positive ($t > 0$) or negative ($t < 0$) risk adjusted return in the data. We will reject the null claim that the risk adjusted return is zero. If $p - value \geq 5\%$, we fail to reject the null hypothesis. That is the data do not provide convincing evidence that the risk adjusted return is positive or negative.

Caution: Dont carelessly use the p-value from regression output. The last column in regression output often lists p-values for one particular hypothesis:

a two-sided test where the null value is zero. If your test is one-sided and the point estimate is in the direction of H_A, then you can halve the softwares p-value to get the one-tail area. If neither of these scenarios match your hypothesis test, be cautious about using the software output to obtain the p-value.

```
# Step 9: Hypothesis Test: t-stats
## Step 9.1: Is the adjusted returns zero
#s1: According to the null
  testValue<-0
  n<-length(marketEXERT)
#s2: compute test statistics
  estcoeff<-Modelsum$coefficients[1]
  eststd<-Modelsum$coefficients[1,2]
  tstats<-(estcoeff-testValue)/eststd
  #or
  tstats<-Modelsum$coefficients[1,3]
#s3: decision rule for two sided test
  decRule<-abs(tstats)>qt(1-0.05/2, n-1-1)
#s4: conclusion
  Result<-ifelse(decRule, "Reject", "Can’t Reject")
```

t Distribution

pt(q, df=n-k-1) is the t probability distribution function with degree of freedom equal to $n-k-1$. n is the sample size and k is the number of independent variables. **q** is a vector of quantiles. **df** is the degrees of freedom. This function can output the percentage of all elements within the t distribution that are lower than the value **q**.
qt(p, df=n-k-1) is quantile function for the t distribution with degree of freedom equal to $n-k-1$. n is the sample size and k is the number of independent variables. **p** is a vector of probabilities.**df** is the degree of freedom. This function can output the cutoff point that **p** of the elements within t distribution are lower than cutoff point.

ifelse(test, yes, no)

It returns a value with the same shape as test which is filled with elements selected from either yes or no depending on whether the element of test is TRUE or FALSE. **test** is an object which can be coerced to logical mode. **yes** return values for true elements of test. **no** return values for false elements of test.

Step 9.2: HT Is the risk exposure higher than one? This slope is only the estimate of the parameter values β_i. We might wonder, is this convincing evidence that the true linear model has a slope higher than one? That is, do the data provide strong evidence that the stock i has market beta larger than one? We can frame this investigation into a one-sided statistical hypothesis test:

$$H_0 : \quad \beta_i \leq 1. \quad \text{The sensitivity is one.}$$
$$H_A : \quad \beta_i > 1. \quad \text{The sensitivity is higher than one.}$$

We would reject H_0 in favor of H_A if the data provide strong evidence that beta exposure to the market is larger than one. To assess the hypotheses, we identify a standard error for the estimate b_i, compute an appropriate test statistic.

In the hypotheses we consider, the null value for the slope is equal to or less than 1, so we can compute the test statistic using the t (or z) score formula:

$$t - stats = \frac{\text{estimate} - \quad \text{null value}}{\text{Standard Error}} \tag{3.10}$$

We can identify the t-test statistic from the software output, shown in the second row and third column (t value). Most of the time, we set the level of significance to be 5%. If $t - stats > t_{df,5\%}$, there is an evidence that the exposure to market risk is higher than one. We will reject the null claim that the risk exposure is one. If $t - stats < t_{df,5\%}$, we fail to reject the null hypothesis. That is the data do not provide convincing evidence that the risk exposure is higher than one.

```
## Step 9.2: Is the sensitivity higher than one?
#s1: According to the null
  testValue<-1
#s2: compute test statistics
  estcoeff<-Modelsum$coefficients[2]
  eststd<-Modelsum$coefficients[2,2]
  tstats<-(estcoeff-testValue)/eststd
#s3: decision rule for two sided test
  decRule<-tstats>qt(1-0.05, n-1-1)
#s4: conclusion
  Result<-ifelse(decRule, "Reject", "Can't Reject")
```